Lab 5

Jenna, Rowan, Ceili

# Rules

In groups of 2 to 4 (self-organized), complete the following.

# Purpose:

This lab will involve using generative AI to potentially improve the introduction and to develop the SOI that would wrap up the report as well as to more easily generate citations to use in writing reports.

# Learning Objectives

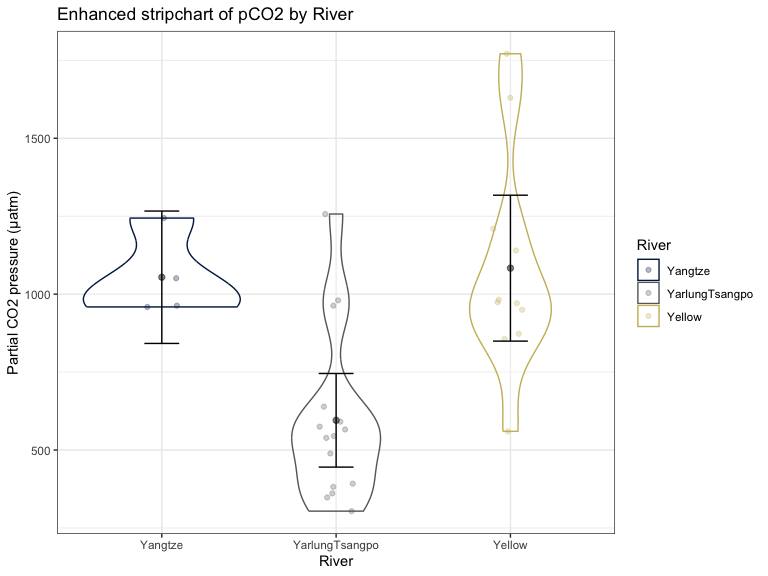
1. Verify that all students have the correct version of the R package ggResidpanel for use to complete the following homework.
2. How to use generative AI to edit and improve writing, but also learn to use those suggested edits with caution.
3. How to write an SOI for a given study using generative AI by providing aspects of the study and then edit/modify results to match the study design.
4. How to use generative AI to aid in developing a references section for a report.

# Task 1:

The following code reads in the data set discussed in the provided “Demonstration Report” and removes observations from the Indus river and makes Figure 1 discussed in the report.

library(readxl)  
riverdata <- read\_excel("Qu\_etal\_2017.xlsx")  
riverdataR <- riverdata %>% dplyr::filter(River != "Indus")

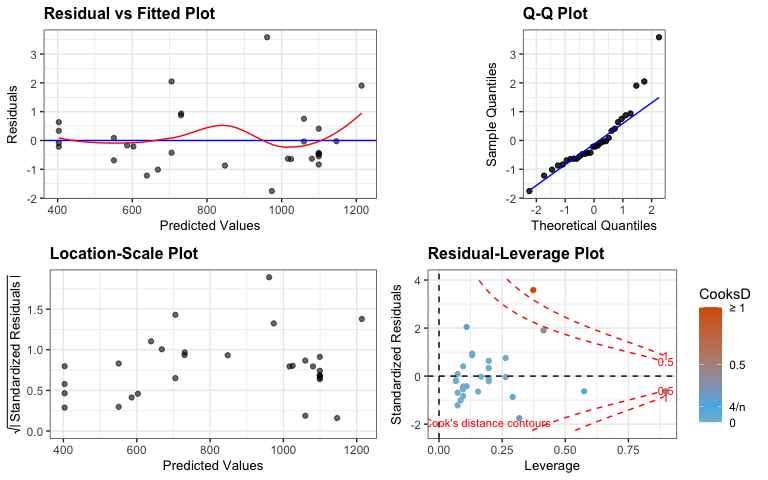
enhanced\_stripchart(pCO2 ~ River, data = riverdataR) + labs(y = "Partial CO2 pressure (\u03BCatm)")



Use the data set to fit the interaction model for the (not the transformed response). Make the standard “R” diagnostic plot using resid\_panel and resid\_interact. Show the results to an instructor to verify the correct version of the package is being used. Using the interactive display, what is the value of the largest Cook’s D for this model. Just report that Cook’s D value - no other discussion.

* **Largest Cook’s D for this model is 1.272.**

lm1 <- lm(pCO2 ~ River \* Elevation, data = riverdataR)  
resid\_panel(lm1, "R")



#resid\_interact(lm1, "R")  
#resid\_calibrate(lm1, plots = c("resid", "ls"), nsim = 4)

* Comment out the resid\_interact line of code to knit the document. **Complete the rest of the assignment via edits to the knitted Word document.**

# Task 2:

Input the following paragraph from the introduction to the “Demonstration Report” into your selected generative AI with a request to modify/edit/improve some aspect or aspects of the provided writing: “Streams play an important role in the emissions of greenhouse gases (GHGs) such as CO2. To better characterize this role, data were collected on three rivers on the Tibetan Plateau: the Yangtze (YZ), the Yarlung-Tsangpo (YT), and the Yellow (YL). Sampling was performed in 2014-2015, collecting partial CO2 pressure (pCO2) at each river site in μatm (Qu et al., 2017) as well as site elevation (in meters) and river name. Since both intra- and inter-river elevations variations may influence atmospheric pressure on dissolved gases, and river characteristics can impact CO2 emissions, we investigated how elevation impacts pCO2 and if those impacts vary across rivers.”

* Report the suggested edits to the paragraph. Then discuss each edit - would you accept them or not, do they suggest other edits you might consider? Do the suggestions match the state of (your) knowledge about the field and the study being considered?
* Optional: Subtask: Ask for modified edits based on a persona (something like: revise those suggested edits as if you were a … working on …) and/or more detailed prompt (make the writing more … or condense the writing or …). Do the suggested edits change? In what way did they change? You can use these modified versions or your original prompt results.
* Optional Subtask: Complete this part of the assignment using different generative AI platforms and compare results. Which platform provided the most useful edits?

Input:

“Streams play a significant role in the emission of greenhouse gases (GHGs), including CO2. To better understand this role, data were collected from three rivers on the Tibetan Plateau: the Yangtze (YZ), the Yarlung-Tsangpo (YT), and the Yellow (YL). Sampling was conducted between 2014 and 2015, with measurements of partial CO2 pressure (pCO2) in μatm at each river site (Qu et al., 2017), as well as site elevation (in meters) and river name. Since both intra- and inter-river variations in elevation may influence the atmospheric pressure on dissolved gases, and river characteristics can affect CO2 emissions, we investigated how elevation influences pCO2 and whether these effects differ across rivers.”

Key changes:

1. “Important role” → “significant role” for slightly stronger phrasing.
2. “Emissions of greenhouse gases” → “emission of greenhouse gases” to improve parallel structure.
3. “Data were collected on three rivers” → “data were collected from three rivers” for more natural phrasing.
4. “Sampling was performed” → “Sampling was conducted” for smoother phrasing.
5. Changed “since both intra- and inter-river elevations variations” to “since both intra- and inter-river variations in elevation” for clearer expression.

ChatGPT mostly suggests small changes in language to be more concise and improve flow. Some of the changes are unnecessary and don’t improve the paragraph in a meaningful way. These changes could be implemented or not and the integrity of the paragraph would remain relatively the same. Each of us asked ChatGPT independently and there were a few edits that were the same for each of us, but it seemed as though even the smallest difference in input wording changed their suggested edits. When prompted to look at changes outside of just grammatical errors it added phrases here and there for clarification, but generally did not change the structure of what we gave it.

# Task 3:

Provide information about the variables and the study design (is there random assignment and, if so, what variables were assigned? and was there random sampling? when and where were the data collected?) and request a sentence that addresses generalizability of the results (can you make inferences to a larger population or not?) and then request a second sentence that addresses whether causal inference is possible in this situation based on the similarly provided context of this study. Make sure you focus on a particular model for the discussion, as the causal aspects might change based on the variables being considered and how they are being used in the model you are focused on.

* Report the suggested generative AI sentences and then edit/modify them to be versions you would want to include at the end of the report in the SOI section. Discuss how you had to edit the results.

Input:

“Ok chat here is the situation: Say I wrote this paper “Greenhouse Gas Emissions in the Rivers of the Tibetan Plateau” by Qu. Here is their sampling methodology and a bit of background about some assumptions that may have been broken. I am working off an interaction model that looks at only pCO2 ~ River\*Elevation. Can you give me 1. A sentence that addresses generalizability of the results (can you make inferences to a larger population or not?) and 2. A sentence that addresses whether causal inference is possible in this situation based on the similarly provided context of this study.” (Added sampling methodology from the paper)

**First Output:**

*Generalizability of the results: The study's findings may be limited in terms of generalizability due to the specific geographic focus on rivers of the Tibetan Plateau and the unique hydrological and meteorological conditions of this region, which may not fully represent other river systems globally.*

*Causal Inference: Causal inference is unlikely to be drawn from this study's observational design, as the analysis focuses on correlations between pCO2 and factors such as river elevation, without controlling for all potential confounding variables or establishing a clear cause-and-effect relationship.*

**We then asked it to keep in mind multiple sampling per site:**

***Generalizability of Results:*** *The study's findings may be limited in terms of generalizability due to potential biases in sampling, as the observation sites were not randomly selected. The selection of more easily accessible sites could introduce systematic differences in pCO2 values, which may not be representative of the broader river population, especially given the variation in geographical distance between sites and the potential clustering of similar sites.*

***Causal Inference:*** *Causal inference is compromised in this study due to potential violations of the independence assumption. The repeated measures within each river and the temporal sequencing of samples could result in observations that are not independent, leading to biased results. Additionally, without controlling for these dependencies or including river as a random effect in the model, the study's conclusions regarding causal relationships between pCO2 and elevation may be unreliable.*

Looking at these two examples that ChatGPT provided, I would work predominantly off of the second options as it gives more context to violations of assumptions. When editing these, I would remove parts such as “Causal inference is compromised”. I would end up with something like the following.

“Findings cannot be generalized widely to other rivers in the area due to sampling biases, with non-randomly selected, easily accessible sites potentially skewing pCO2 values. Non-random selection of these four river systems limits applicability to other river systems globally.Conclusions regarding the relationship between pCO2 and elevation are biased due to violations of the independence assumption, as repeated measures within each river and temporal sequencing of samples result in non-independent observations. Without random assignment of elevation and river systems we cannot infer causal inference.”

Edited further we may end up with something like the following:

“The results of this study cannot be generalized to a wider population of global high-elevation river systems due to potential sampling bias as the observation sites were not randomly sampled. Sites were selected by ease of access which could cause a misrepresentation of the pCO2 values. As this study is observational, there is no random assignment of elevation or river and we cannot draw causal inferences. The findings can be generalized to the sample of rivers in the Tibetan Plateau.”

# Task 4:

Formatting citations can often be challenging when they come from disparate sources. Extract the noted citation information from various locations in the provided report using citation(“Rpackagename”) in R (console is fine - no need to go back to the provided Rmd) for obtaining citation information for the R packages used and other sources for other citations. Note that all citations were highlighted in the demonstration report to help you to identify the needed citations. Provide the information to your selected generative AI and use it to make sure the formatting is consistent with a particular citation style and in alphabetical order.

* Submit your citations/references section for feedback here. Eventually this will be a question on the next HW.

Arel-Bundock, V. (2022). "modelsummary: Data and Model Summaries in R." Journal of

Statistical Software, 103(1), 1-23. <https://doi.org/10.18637/jss.v103.i01>.

Arnold, J. (2024). ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R package

version 5.1.0. <https://CRAN.R-project.org/package=ggthemes>.

Fox, J., & Weisberg, S. (2019). An R Companion to Applied Regression, Third edition. Sage,

Thousand Oaks, CA. <https://www.john-fox.ca/Companion/>.

Goode, K., Rey, K., & Greenwood, M. (2025). ggResidpanel: Panels and Interactive

Versions of Diagnostic Plots using 'ggplot2'. R package version 0.3.1, commit

bbabb7cab3ea0501fe8af65b7cd4c7477d7088d8. [https://github.com/greenwood- stat/ggResidpanel](https://github.com/greenwood-%20stat/ggResidpanel).

Greenwood, M. (2025). catstats2: Upper Level Statistics for Montana State University

Bobcats. R package version 0.2, commit 02be2a5c90da261042a329833ca15de19be6cf9c.

<https://github.com/greenwood-stat/catstats2>.

Greenwood, M., Hancock, S., & Carnegie, N. (2024). catstats: Statistics for Montana State

University Bobcats. R package version 0.1, commit

f6b0caa7074a18e9ed2669efd754dae29c03be45. [https://github.com/greenwood- stat/catstats](https://github.com/greenwood-%20stat/catstats).

Jeppson, H., Hofmann, H., & Cook, D. (2021). ggmosaic: Mosaic Plots in the 'ggplot2'

Framework. R package version 0.3.3. <https://CRAN.R-project.org/package=ggmosaic>.

John Fox, & Weisberg, S. (2019). An R Companion to Applied Regression, 3rd Edition.

Thousand Oaks, CA.

<https://socialsciences.mcmaster.ca/jfox/Books/Companion/index.html>.

Koneswarakantha, B. (2023). easyalluvial: Generate Alluvial Plots with a Single Line of

Code. R package version 0.3.2. <https://CRAN.R-project.org/package=easyalluvial>.

Lenth, R. (2024). emmeans: Estimated Marginal Means, aka Least-Squares Means. R

package version 1.10.6. <https://CRAN.R-project.org/package=emmeans>.

Pedersen, T. (2024). patchwork: The Composer of Plots. R package version 1.3.0.

<https://CRAN.R-project.org/package=patchwork>.

Pruim, R., Kaplan, D. T., & Horton, N. J. (2017). The mosaic Package: Helping Students to

'Think with Data' Using R. The R Journal, 9(1), 77-102.

Sievert, C. (2020). Interactive Web-Based Data Visualization with R, plotly, and shiny.

Chapman and Hall/CRC, Florida.

Sjoberg, D. D., Whiting, K., Curry, M., Lavery, J. A., & Larmarange, J. (2021).

Reproducible summary tables with the gtsummary package. The R Journal, 13, 570–80.

<https://doi.org/10.32614/RJ-2021-053>.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New

York.

Wickham, H., & Bryan, J. (2023). readxl: Read Excel Files. R package version 1.4.3.

<https://CRAN.R-project.org/package=readxl>.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R.,

Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E.,

Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K.,

Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse.

Journal of Open Source Software, 4(43), 1686. <https://doi.org/10.21105/joss.01686>.

# Task 5:

Reflect on the use of generative AI for each of the three tasks. How well did your chosen generative AI do the requested tasks? Rank your chosen generative AI in its usefulness on the three previous tasks and explain the reason for your ranking.

If we were to rank the tasks for usefulness we would place editing of citations as the most useful. We asked for citations a variety of ways, including giving it an already created list to alphabetize and ensure consistent formatting as well as asking it to create the entire list itself. In both of these scenarios it performed quite well and was even able to find the Catstats2 citation from github. The second most useful output it gave us was dealing with grammatical proofreading. Since this part of proofreading is much more mathematical, generative AI did a good job at assisting it “ironing out” some of the writing. That being said, we did not always agree on the changes that it made, especially when it would suggest word changes like including “significant” for statistical writing. Where generative AI showed the least usefulness was in the creation of its own sentences, such as scope of inference and conclusions. To get the output to where you want it, in your particular voice, you need to provide a lot of output. At that point we felt as though it would be easier to simply take the time to write it yourself.

# Task 6:

Moving forward, will you use generative AI for similar tasks in the future? What ways have you or do you expect to use generative AI in your academic career?

Generative AI is something that we think we may use in the future for citation formatting. Other models like Zotero do a wonderful job at creating reference lists for long papers, but we have noted a consistent difference in formatting for different sources. It may be useful to plug the citation list then into ChatGPT to ensure consistent formatting without a whole lot of manual labor. We will not be using use generative AI for writing as the voice is very mechanical and specific information input is picky. That being said, it is good for proofreading, and we may use it for that when it comes to assignments or emails etc where a specific scaffolding is not required.

\*Side note: Ceili in the group works a lot with remote sensing and deep learning models. Much of the work cannot be done at the scale it is without these awesome tools, and a lot of the work comes down to knowing how to communicate with the computer. (They only know what they have been told!) It is never good to take an output at face value, always validate and recognize that there will be errors.

# Grading Rubric:

Each task will be assessed on a scale from 0 to 5, with 0 being not attempted, 1 being minimal effort or incomplete reporting of results, 2 being partially incomplete or poor effort, 3 being a moderate effort but notable missing/incorrect on some parts, 4 being good work on some parts but missing/incomplete on others, and 5 being excellent and complete work.